



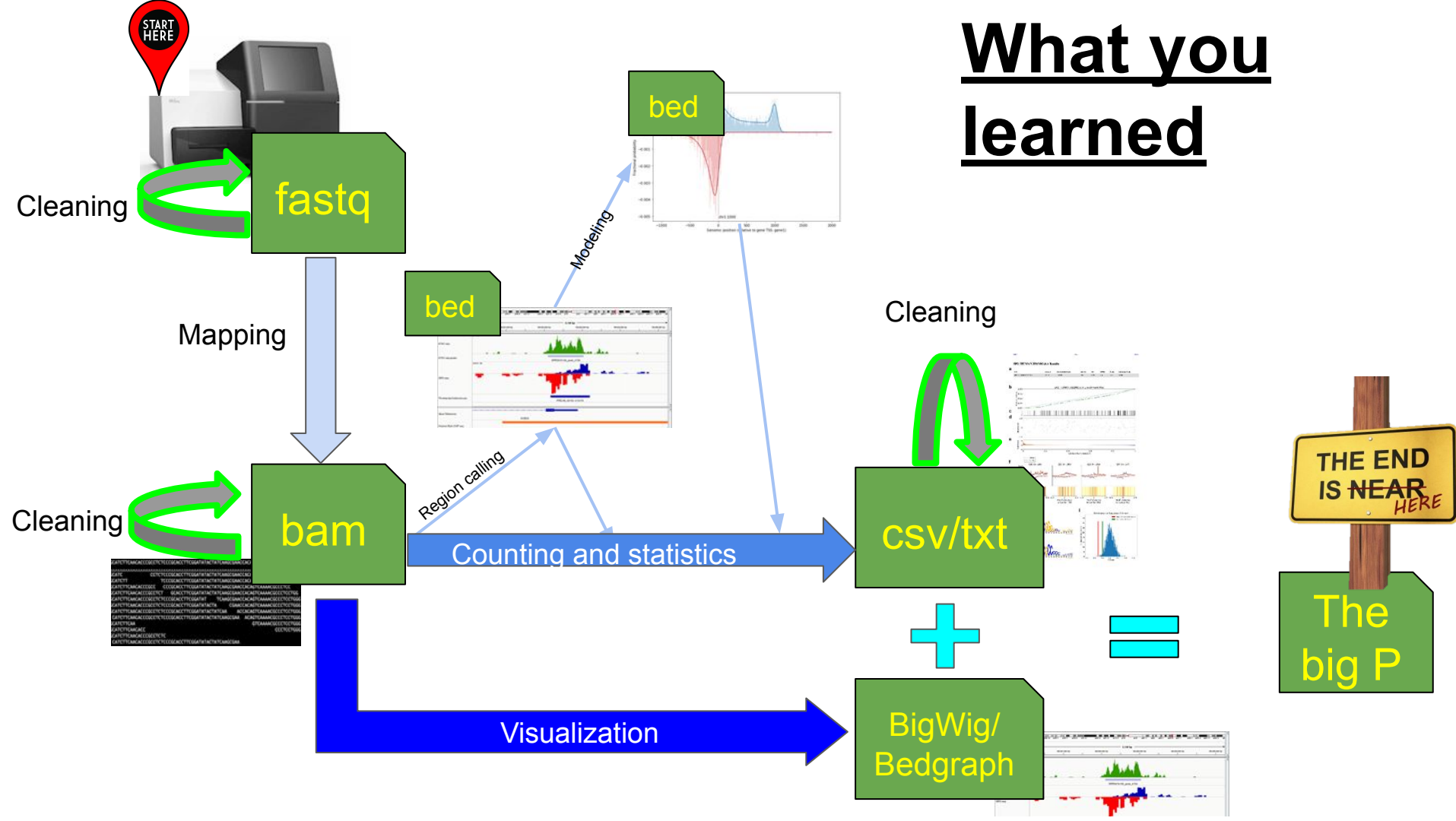
# DnA Lab

## Short read sequencing

## Workshop

Day 10

# What you learned



```
[root@localhost ~]# cd /var
[root@localhost var]# ls -la
total 72
drwxr-xr-x. 18 root root 4096 Jul 30 22:43 .
drwxr-xr-x. 23 root root 4096 Sep 14 20:42 ..
drwxr-xr-x. 2 root root 4096 May 14 00:15 account
drwxr-xr-x. 11 root root 4096 Jul 31 22:26 cache
drwxr-xr-x. 3 root root 4096 May 18 16:03 db
drwxr-xr-x. 3 root root 4096 May 18 16:03 empty
drwxr-xr-x. 2 root root 4096 May 18 16:03 games
drwxrwx--T. 2 root gdm 4096 Jun 2 18:39 gdm
drwxr-xr-x. 38 root root 4096 May 18 16:03 log
drwxr-xr-x. 2 root root 4096 May 18 16:03 logrotate.d
drwxr-xr-x. 2 root root 4096 Jul 1 22:11 repodata
lrwxrwxrwx. 1 root root    6 May 14 00:12 run -> /run
drwxr-xr-x. 14 root root 4096 May 18 16:03 spool
drwxrwxrwt. 4 root root 4096 Sep 12 23:50 tmp
drwxr-xr-x. 2 root root 4096 May 18 16:03 yp
[root@localhost var]# yum search wiki
plugins: langpacks, presto, refresh-packagekit, re
primary_db
73% [=====
```

Start

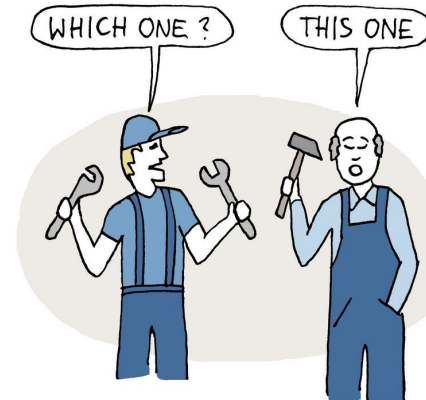


# **SAFETY REMINDER**

**USE THE  
RIGHT TOOL**  
— s s s —  
**FOR THE  
RIGHT JOB**  
— s s s —  
**IN THE  
RIGHT WAY**



creativesafety.com 866-777-1360 #WS30010-24



# Today

- Capstone day 10
- Github, Downloading/Uploading data
- Your questions answered!! (we hope)

## Day 10 homework

Keep practicing -- and keep in touch! We would love to hear from you!

The  
expert in  
anything  
was  
once a  
beginner.



While these terms are often used interchangeably, there is an important nuance. The chef is a trailblazer, the person who invents recipes. She knows the raw ingredients and how to combine them. The cook, who reasons by analogy, uses a recipe. He creates something, perhaps with slight variations, that's already been created.



The difference is real knowledge as opposed to know-how.

## How do I continue my training?

There are numerous online courses in data analysis, data science, programming, machine learning, and visualization.

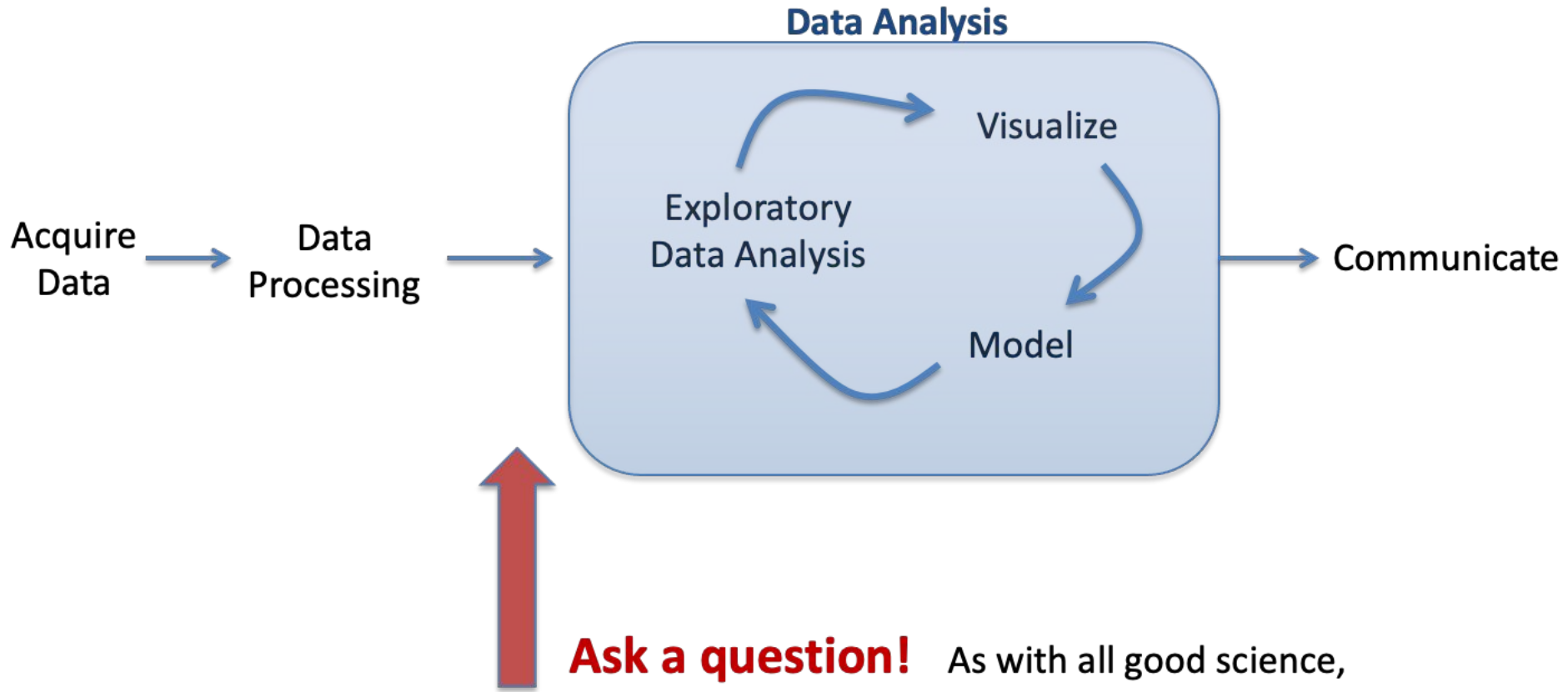
*READ:* Books, online articles

**Just do it** – practice. Find avenues to use your newly acquired skills.



*Practice*  
— —  
**makes**  
— —  
*perfect*





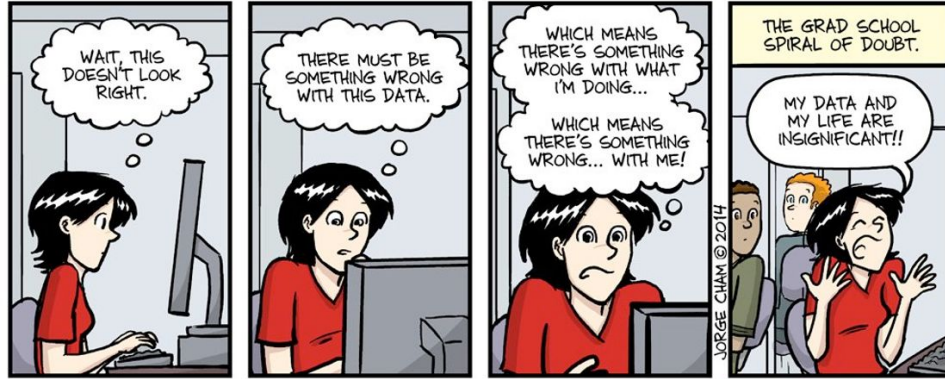
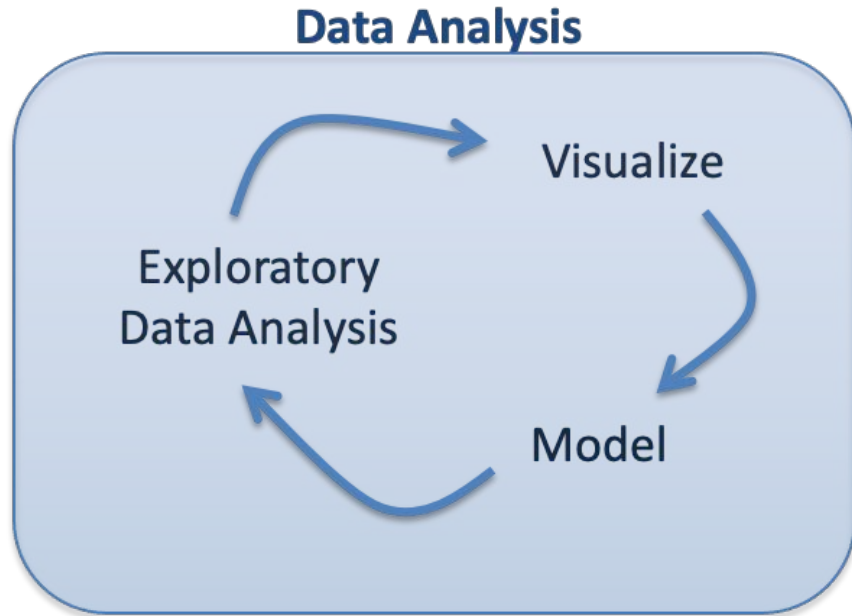
**Ask a question!** As with all good science, analysis of data starts with a question.

“The **data** may not contain the answer.  
The combination of some data and an  
aching desire for an answer does not  
ensure that a reasonable answer can be  
extracted from a given body of data.”

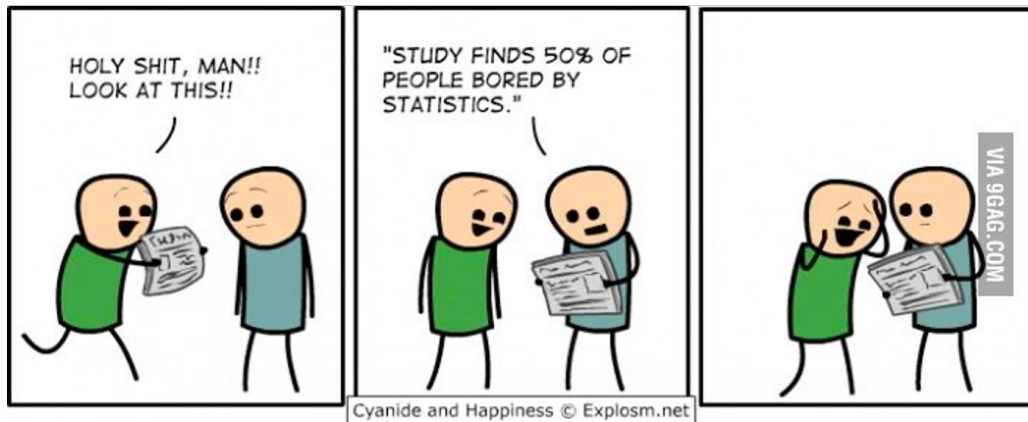
---John Tukey

Quality control helps, but doesn't address this completely.

# Inside the analysis box ...



# A few words on statistics ...



“It’s easy to lie with statistics, but it’s hard to tell the truth without them.”

— Charles Wheelan

FOURTH EDITION

# INTUITIVE BIOSTATISTICS

A Nonmathematical Guide to Statistical Thinking

Harvey Motulsky

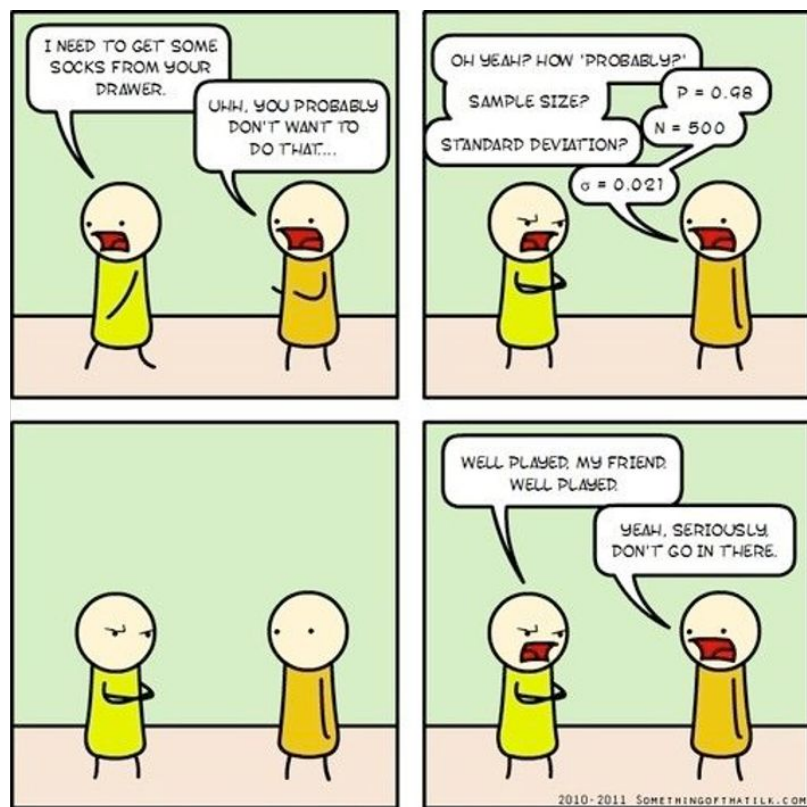
OXFORD  
UNIVERSITY PRESS

Copyrighted Material

O'REILLY

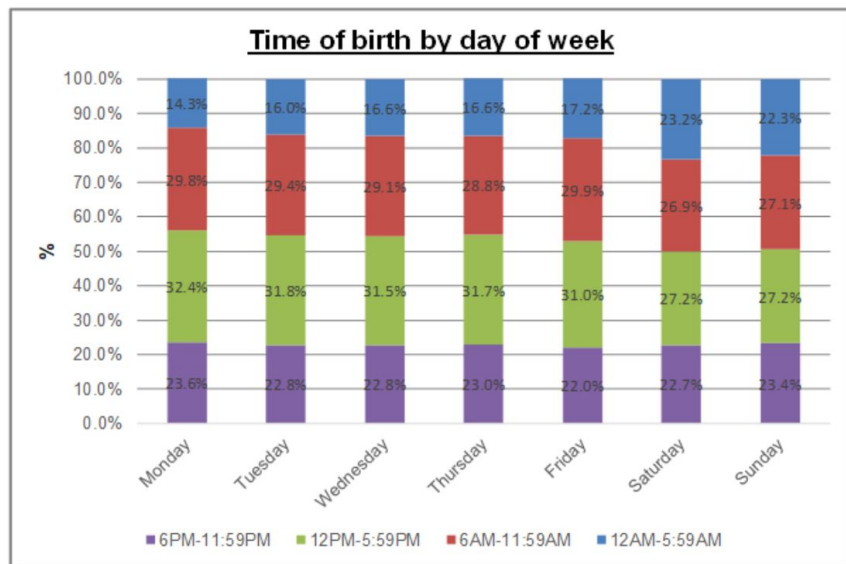
## Practical Statistics for Data Scientists

50 ESSENTIAL CONCEPTS



# Communicate clearly your data and analysis ..

Ugly graphs are useful for rapid visualization:

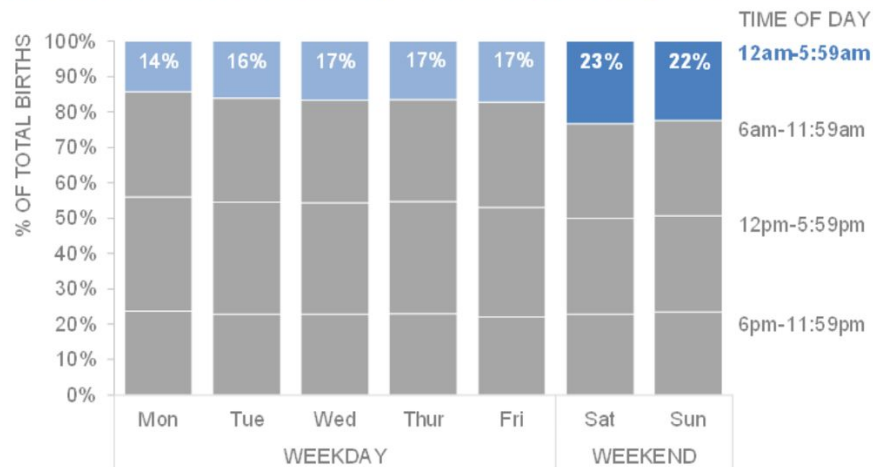


Data source: CDC (National Vital Statistics Reports, Vol. 67, No. 1, January 31, 2018)

But publications and talks call for pretty, message driven graphs:

When babies are born

Weekend deliveries are more likely to be in early morning, compared to weekdays



Data source: CDC (National Vital Statistics Reports, Vol. 67, No. 1, January 31, 2018)

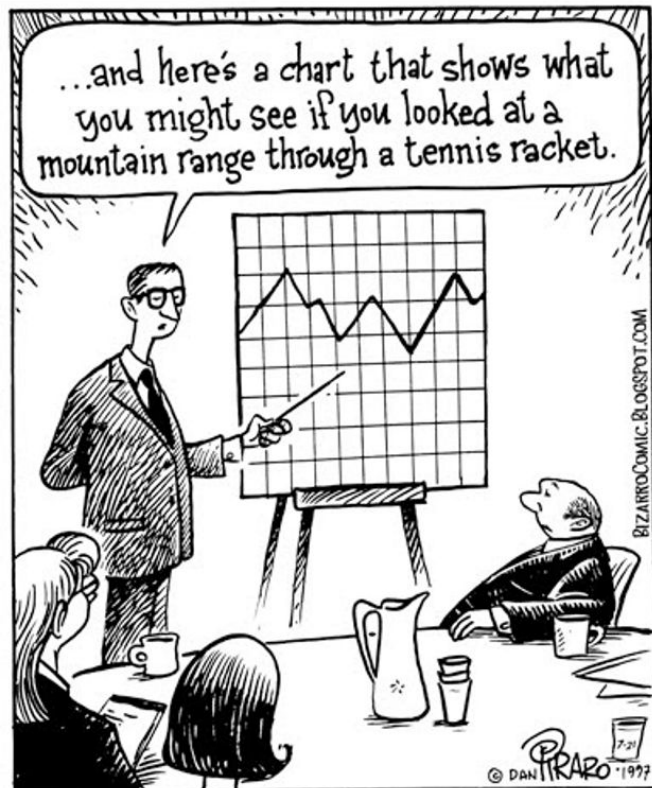


cole nussbaumer knaflic

# storytelling with data

a data  
visualization  
guide for  
business  
professionals

WILEY

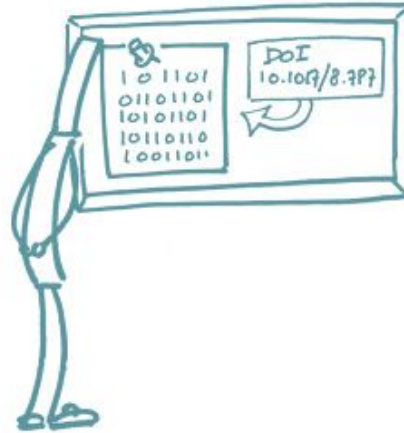


# YOUR data should adhere to FAIR principles

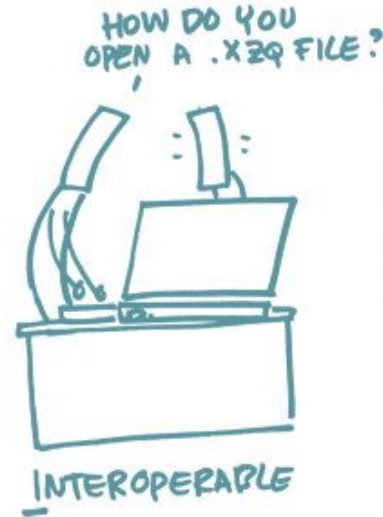
## FAIR DATA PRINCIPLES



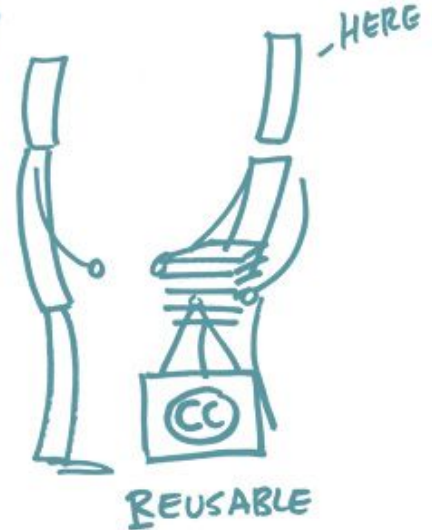
FINDABLE



ACCESIBLE



INTEROPERABLE



REUSABLE

# One final word: Skepticism is a healthy skill to have ...



ELSEVIER

## The American Journal of Medicine

Available online 2 December 2022

In Press, Corrected Proof ?



Clinical Research Study

## COVID Vaccine Hesitancy and Risk of a Traffic Crash

Donald A. Redelmeier MD, FRCPC, MSHSR, FACP <sup>a, b, c, d, e</sup>  , Jonathan Wang MMASc <sup>b, c</sup>,  
Deva Thiruchelvam MSc <sup>a, c</sup>

Show more ▾



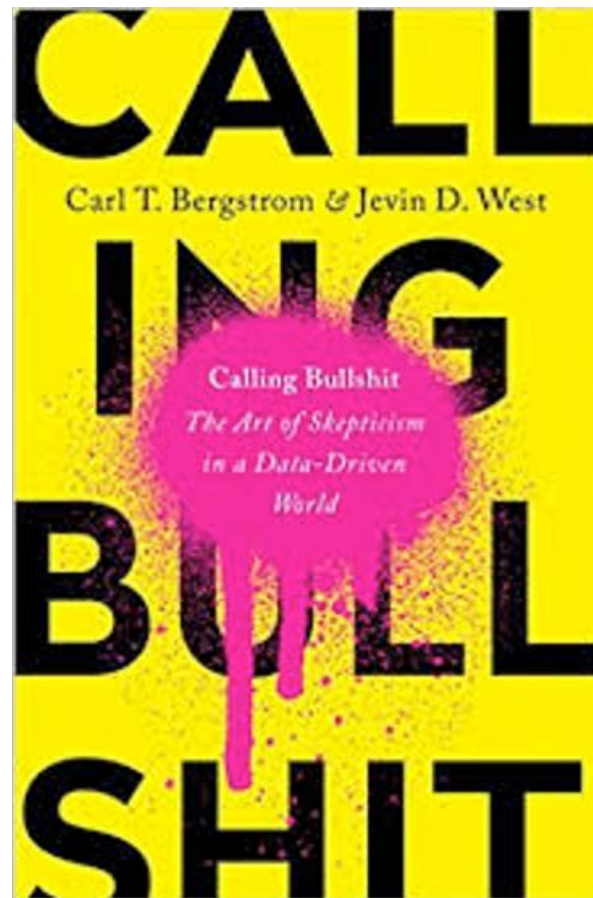
How to Steer Clear of Common Blunders When Working  
with Data and Presenting Analysis and Visualizations

# AVOIDING DATA PITFALLS

BEN JONES

Founder and CEO, Data Library

WILEY



<https://www.callingbullshit.org/>

The Website will stay up!  
The github will stay up!

The AWS supercomputer will die on Monday.